

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-241227

(43)Date of publication of application : 17.09.1996

(51)Int.Cl.

G06F 11/34

G06F 3/06

(21)Application number : 07-328336

(71)Applicant : HEWLETT PACKARD CO <HP>

(22)Date of filing : 22.11.1995

(72)Inventor :
VOIGT DOUGLAS L
NELSON MARVIN D
DOLPHIN WADE A

(30)Priority

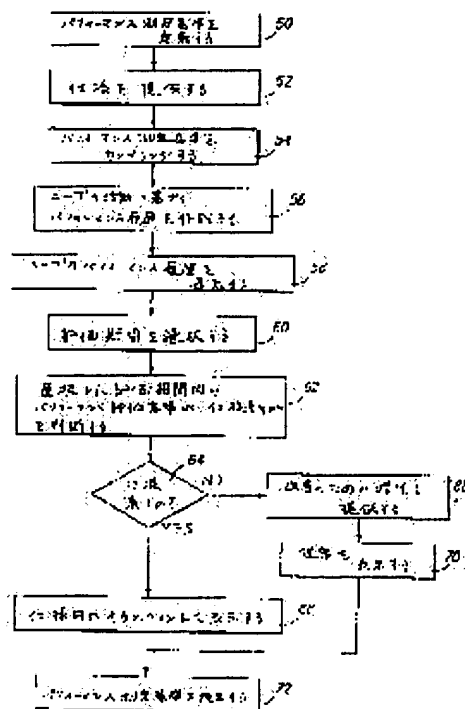
Priority number : 94 343809 Priority date : 22.11.1994 Priority country : US

(54) METHOD FOR IMPROVING AND SPECIFYING PERFORMANCE OF STORAGE SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a method for improving and specifying performance of storage system by which the performance of an operated storage system is evaluated with the purpose of monitoring or predicting such a state that gives an adverse influence on the performance and, when it is detected that the performance is not optimum, the user of the system is warned of it so that the user can take a measure to improve the performance.

SOLUTION: A performance measuring standard is defined for measuring a plurality of performances during the operation of a data storage system (5) and a specification is provided for the expected value (52). During the operation of the system, the performance measuring standard is sampled (56) and a performance hysteresis is generated and how the system operates during a fixed period is displayed (58). An evaluating period is selected and whether or not an arbitrary measuring standard meets the specification is discriminated (62). When one or more performance-measuring standards do not meet the specification, one or more proposals which are useful for improving the performance of the data storage system are given (68).



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2000 Japan Patent Office

* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] The step which is the technique of pinpointing the field for the enhancement in performance in a data-storage system, and samples performance metrics working [data storage], The step which shows an user the performance history which shows how the data-storage system is operating during fixed, The step as which an user can choose the evaluation term of a performance history, The step which judges whether performance metrics have the level of the optimum performance of a data-storage system during the selected evaluation, How to have the step which provides an user with at least one proposal considered to be useful to raising the performance of a data-storage system when there are no performance metrics in the level of optimum performance.

[Claim 2] Technique according to claim 1 that a presentation step is characterized by an user having the step which enables it to choose the evaluation term of a performance history by enabling it to operate the term field where the selection step piled up the performance history on the performance history on the display screen including the step which carries out graphical display by the graphics.

[Claim 3] Technique according to claim 1 that a decision step is characterized by having the step which takes the average of the sample of the performance metrics measured during fixed, and asks for type average performance metrics.

[Claim 4] Technique according to claim 1 characterized by having the step with which a decision step combines the sample of the performance metrics measured during fixed, and the step which calculates the type value which filters the combined sample and shows performance metrics.

[Claim 5] The step which defines two or more performance metrics which are the technique of pinpointing the field for the enhancement in performance in a data-storage system, and can be measured working [a data-storage system], The step which a data-storage system presupposes is operated with optimum performance when the specification of performance metrics is offered and performance metrics follow the specification, The step which samples performance metrics working a data-storage system], The step which creates the performance history which shows how the data-storage system is operating during fixed based on the storage demand which the user issued, The step which shows an user a performance history, and the step as which an user enables it to choose the evaluation term of a performance history, The step which judges whether there are any performance metrics according to a specification during the selected evaluation, When one or more performance metrics do not follow a specification Technique characterized by having the step which offers the proposal considered to be useful to raising the performance of a data-storage system based on any of one or more performance metrics do not follow the specification.

[Claim 6] Technique according to claim 5 that it has the threshold of each performance metrics, and a decision step takes the average of each performance metrics in an evaluation term, and a specification is characterized by having a step [metrics / mean performance / each threshold].

[Claim 7] Technique according to claim 5 that a decision step is characterized by having the step which checks how much one or more performance metrics do not follow the specification.

[Claim 8] Technique according to claim 5 characterized by for a specification having two or more thresholds of each performance criteria, and having the step to which the decision step of each performance metrics takes the average of the performance metrics in an evaluation term, a step [metrics / mean performance / two or more thresholds], and the step which computes the grade of a nonconformity based on any of two or more thresholds are over performance metrics.

[Claim 9] Technique according to claim 5 that it has the step to which a presentation step carries out graphical display of the performance history on the display screen, and a selection step piles up a term field by the graphic on a performance history, and is characterized by having the step which enables it to choose an evaluation term when an user operates a term field by the graphic.

[Claim 10] Technique according to claim 5 that it has the step which lists two or more predetermined proposals on the table which relates a proposal with performance metrics, and a proposal offer step is characterized by having the step which extracts one or more proposals from a table based on any of one or more performance metrics are not based on the specification.

[Translation done.]

* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention] this invention relates to data-storage systems, such as a disk array and a data-storage system, and the method of specifying the enhancement technique in performance based on a demand of an user.

[0002]

[Description of the Prior Art] The performance of a data-storage system changes inevitably into the service life. Performance is influenced of various factors, such as a type of a component used by the storage system, the set point of operation, capacity, and requirements for storage of an user like input/output request. When a component causes failure or an activity ratio becomes too much or inadequate, performance may receive a bad influence. When the set point of operation is unsuitable for an user's storage configuration, performance may fall. When an user issues the storage demand which is too severe for a present system, too much load will be applied to a storage system, and performance will receive a bad influence. [0003] Many factors which affect such performance change into the service life of a data-storage system. Furthermore, such a factor may be changed sharply for a short period of time. When performance changes remarkably, if an user raises performance, he is often going to add change to the system which it predicts or is sure of. However, although such change is performed based on the ****ed problem, the problem may be the indication of the actual more fundamental hidden problem. By the problem solving which is not based on such an information, an user may produce an unnecessary cost and system stop time, without obtaining the satisfactory enhancement in performance. For example, although performance improves only by reconfiguring a data storage by the new method corresponding to the increase in a storage demand of an user in fact, if an user adds one more disk to a disk array, it may be thought that performance improves.

[0004]

[Problem(s) to be Solved by the Invention] Therefore, it is advantageous **** if a working system and performance are evaluated in order to supervise or predict the status which affects performance. If it is detected that the performance of a system is not the optimum, it can warn an user of it and can direct about the treatment which can lecture in order to raise performance.

[0005]

[Means for Solving the Problem] According to one mode of this invention, the method of pinpointing the field which should improve in a data-storage system is offered. Two or more performance metrics which can be measured working [a data-storage system] are defined. As an example of performance metrics, there is the length of the time when the command was published etc. from the sum block count which the host read, the sum block count which the host wrote in, and a host. The specification of the expected value of performance metrics is also offered. When performance metrics follow the specification, the data-storage system is operating with optimum performance.

[0006] According to this technique, performance metrics are sampled working [a data-storage system]. Working, a performance history is created and it is shown how the data-storage system was operating during fixed. By the performance history, the activity and storage demand which were performed by users, such as the number of I/O per second, the number of memory transactions, and a used and intact room, are inspected. As for a performance history, it is desirable that it is the type of the graph which an user is provided with and plotted the activity and storage demand during a fixed term.

[0007] When investigating a performance history, an user can choose an evaluation term in order to scrutinize to a precision more. As for this, realizing, as an evaluation term can be chosen is desirable by piling up a term field on a performance history using a graphic, and an user's operating the term field by the graphic, and putting it on a desired time frame. It can judge whether about the selected evaluation term, arbitrary performance metrics follow the specification. When one or more performance metrics do not follow a specification, one or more various proposals which are useful to raising the performance of a data-storage system are given. It is based which performance metrics do not follow the specification and a proposal is chosen.

[0008] By use of these proposals, an user can determine the technique of raising the performance of a data-storage system based on an information. Moreover, by the performance plan technique of this invention, an user can predict the required storage capacity of the future, and he can also upgrade a data-storage system in timely so that it may correspond to it.

[0009]

[Example] The technique of this invention is explained about the computer disk array and data-storage system which uses two or more disk drives. However, the principle of this invention is applicable also to the data-storage systems of other types, such as an open-reel storage system, and tape backup system, CD storage system.

[0010] The computer system 10 which has the host-computer terminal or the station 12 connected to drawing 1 through the input/output interface and the bus 16 at data storage 14 is shown. The host-computer station 12 is equipped with the display monitor 18, the arithmetic and program control (CPU) 20, and the keyboard 22.

[0011] The example of a configuration of the data-storage system 14 carried out as a hierarchy disk array and a data-storage system is shown in drawing 2. It connects with the disk array 30 and the disk array 30 which have two or more storage disks 32, and the storage system is equipped with the disk array control unit 34 which adjusts the data transmission between the storage disks 32, and the RAID managerial system 36.

[0012] A part of physical storage capacity is used for the vocabulary "RAID" (redundant array of an independent disk), and it means the disk array which memorizes the redundant information about the user data memorized by the remaining fraction of storage capacity. When failure occurs on one or the access pass of the configuration disk of an array, regeneration of the user data can be carried out by this redundant information. " It is The RAIDBook. about the detail of RAID system, it was published from RAID Advisory Board of the American Minnesota ***** on June 9, 1993. : It is indicated by book called A Source Book for RAID Technology."

[0013] The disk array control unit 34 is connected to the disk array 30 through one or more interface buses 38, such as a small computer and the SI (SCSI). Through the interface protocol 40, as the RAID managerial system 36 exerts an operation on the disk array control unit 34, it is connected to it. The control unit 34 is connected to the host computer through the input/output interface and the bus 16. The RAID managerial system 36 can be carried out as a separate component, or can be constituted in the disk array control unit 34 or a host computer, a disk memory and reliability level are controlled, and the data control means for transmitting data among various reliability storage level is offered. Although it is desirable that they are a mirror or parity redundancy level as for such reliability storage level, it can also include the reliability storage level without a redundancy.

[0014] The disk array control unit 34 is carried out as dual-control equipment equipped with disk array control unit A 34a and disk array control unit B34b. The dual-control equipments 34a and 34b strengthen a reliability by offering the backup and redundancy without a way piece, when one control unit becomes the impotentia of operation.

[0015] The storage disk 32 has 1 or the capacity of 3 gigabytes. It can connect independently with the machine bay which carries out an interface to the SCSI bus 38, respectively, or a storage disk can be separated. In the one embodiment, the data-storage system equipped with 12 activity machine bays is designed. Four SCSI buses can be used and the interface of these bays and the disk array control unit 34 can be carried out (that is, one bus is used about three machine bays). If all activity bays are loaded, the compound capacity of instantiation of this data-storage system will become 12 or 36 gigabytes.

[0016] The hierarchy disk array 30 can be characterized as a respectively separate storage space containing the physical storage space and one or more virtual-memory space. These various views of storage are associated by mapping technique. For example, the map of the physical storage space of a disk array can be carried out to the virtual-memory space of RAID level which expresses a storage area according to various data-reliability level. For example, the field of the part in the virtual-memory space of RAID level can be assigned because of the 1st reliability storage level, such as a mirror or the RAID level 1, and other fields can be assigned because of the 2nd reliability storage level, such as parity or the RAID level 5. The map of the virtual view of RAID level can be carried out to the virtual-memory space of the 2nd application level which shows the storage space which can be consecutive-addresses specified. The physical configuration and RAID view of a storage space are hidden from the application view shown to an user.

[0017] The memory map storage 42 is prepared for the self-sustaining storage of a virtual mapping information used in order to carry out the map of various storage spaces mutually. A memory-mapping information does not cease with a control unit or RAID managerial system, or can be periodically updated as various mapping configurations between various views change strangely. The memory map storage 42 is carried out as two nonvolatile RAM (NVRAM) 42a and 42b, such as RAM (random access memory) by which the battery back-up was carried out, and is arranged in each control unit 34a and 34b. Double NVRAM 42a and 42b is formed for memory-mapping information-redundancy storage.

[0018] A storage disk offers the physical storage capacity which a host computer can use for a data storage. It is desirable to manage this capacity effectively and to optimize performance from the viewpoint of both that a storage demand of luminous efficacy of operation and an user is satisfactory. The footing of the performance management technique of this invention has a data-storage system in finding out the reply to the difficult and inaccurate question of evaluating continuously which operating effectively. The question of the following seven term shows the example of 1 set of enhancement decision-making targets in performance.

[0019] 1. Is the data-storage system interference of performance?

2. Is optimum performance possible by the present system configuration set point? As an example of the set point, there are SCSI transfer rate, a size of a logical unit, etc.

Does performance improve by restarting the balance of the storage disk connected to 3.4 SCSI channels 38?

4. a storage disk -- large-capacity-izing -- if it accelerates or adds, does performance improve sharply?

5. If RAM in the disk array control unit 34 is added, does performance improve sharply?

6. If the data-storage system 14 is added, does performance improve sharply?

7. If it upgrades to the disk array control unit 34 or the storage disk 32 of a new generation, does performance improve?

[0020] Even if a cause is not traced, something shows having done the bad influence to the system, and the question of decision making mentioned to the 1st term shows that the question 2 or five term needs to add change to the existing data-storage system. The question 6 and seven term require change which needs an additional component, such as change after all using up the capacity of one data-storage system.

[0021] Since a certain amount of forecast is needed, it is difficult to reply to the question about the enhancement in performance with absolute accuracy. However, the high storage cost and the analysis processing cost are required for the perfect quantity-estimate of a performance parameter. Therefore, a focus is doubled with the existing component of a data-storage system being used for the technique of this invention. Operating-condition data are useful to specifying the problem of performance, and are useful to specifying the component which is operating effectively with the present configuration, and an inferior component. The enhancement strategy in basic performance after specialization is not at change of a component which has already operated by specification within the limits, but is making an investment [change / which solves the present problem].

[0022] The technique of this invention pinpoints the potential field for the enhancement in performance. Such technique tells an user about the performance below a canonical, and generates the proposal which supports that an user corrects a problem and returns a system to optimum performance.

[0023] The desirable step of the technique of pinpointing the field for the enhancement in performance is shown in drawing 3 . The first step 50 is defining one or more performance metrics which can be measured working [a data-storage system]. As an example of a performance criteria, there are the length of the time when the command was sent, SCSI queue length per

equipment, etc. from the sum block count which a host reads, the sum block count which a host writes in, and a host. The specification which sets up the canonical of performance metrics is also offered (step 52). The data-storage system is operating with optimum performance, when performance metrics follow this specification, it puts in another way and performance metrics are "as a specification." The canonical set up with a specification can take the gestalt of the acceptance calculation numeric value by the threshold of various performance metrics, or a permission calculation numeric value.

[0024] After defining performance metrics and the specification corresponding to it, a means to supervise performance metrics to a data-storage system is established. As for such a means, it is desirable to carry out as a firmware in the disk array control unit 34 which extracts working component operating-condition data.

[0025] After introducing performance metrics, the following step 54 is sampling the working performance metrics of data storage. This sampling is performed at the data-storage system 14 or the host-computer station 12. Example of ** type 80 of a control unit firmware which showed the point which can sample performance metrics in drawing 4 is shown. A circle expresses storage of radial transfer and the box with three sides expresses a queue. This ** type view shows the information flow from a host computer to a storage disk. The demand from a host is SCSI of a front end first. It flows to 82 and, next, is temporarily put in by a host's queue 84. For processing after it, a decoder 86 subdivides a demand to a parvus command more. It is put into a command by the cache queue 88 until an acceptance setup of a cache 90 is completed. According to the RAID object 94 about a hierarchy and a redundancy management, a work load is changed by the hierarchy manager 92. Flowing is down-stream SCSI. The item sent to 98 follows the SCSI queue 96 into which it is put temporarily.

[0026] Performance metrics can be extracted at various points on this ** type. For example, the metrics of the sum block count which the host read are front end SCSI. It is extracted by A points of 82. The decision influenced by the example of metrics, the extracting point of performance metrics, and performance metrics is shown in the following table 1.

[0027]

[Table 1]

表 1		
測定基準	サンプリング点	支援された決定
活動時間—ホストから>0命令が顕著な間の時間の合計	A	1) ボトルネックがあるか？ 7) アップグレードでパフォーマンス改善できるか？
SCSI待ち行列最大値—装置当りの最大のSCSI待ち行列サイズ	D	3) ディスクのバランスを取り直すか？ 4) ディスクの追加は助けになるか？
ホスト読取りブロッカー—ホストにより読み取られたブロックの数の合計	B	1) ボトルネックがあるか？ 5) 制御RAMがもっと必要か？ 7) アップグレードでパフォーマンスを改善できるか？
ホスト書き込みブロック数—ホストにより書き込まれたブロックの数の合計	B	1) ボトルネックがあるか？ 5) 制御RAMがもっと必要か？ 7) アップグレードでパフォーマンスを改善できるか？

[0028] Performance metrics can be measured in the arbitrary locations in a data-storage system. For example, in C points, the sample of the performance metrics about the maximum length of a host queue can be extracted.

[0029] It returns to drawing 3 and the performance history based on an user's storage activities is created at step 56. The parameters reflecting a data-storage system use, such as parameters, such as an amount of the amount of usable space, the amount of the consumed space, input/output request, a mirror, and parity type storage, are contained in a performance history. By the performance history, it becomes clear how the data-storage system was operating during fixed. A creation of the performance history in step 56 can be performed at the data-storage system 14 or the host-computer station 12.

[0030] Next, an user is shown a performance history at step 58 at the host-computer station 12. As for a performance history, it is desirable to be displayed on the display screen 18 of drawing 1 in a graph.

[0031] The example 100 of the display screen of a performance history is shown in drawing 5. Specifically, the unused-memory capacity (it measures per ***** of the data-storage system over three days, the storage capacity (it measures per ***** of RAID level 1 type, and the performance about ON **** access activities (it measures with both the number of kilobytes per second and the number of I/O per second) are shown. Please observe that the remarkable activity is shown at the daytime from 6:00 a.m. to 12:00 p.m. for a performance history. An user can adjust a display term.

[0032] Here, suppose that an user wants to evaluate one specific term. For example, an user observes the unusual moment and unusual inclination within a performance history, and has the case where he wants to investigate the term still minutely. Moreover, in another example, while the user is analyzing the comparatively normal term, there is a case where it is made to want to display the term of a non-type-operating condition. In both cases, an user operates the term field divided by the vertical lines 102 and 104 piled up on the performance history graph by the graphic, and just needs to choose an evaluation term (step 60 of drawing 3). Each boundary lines 102 and 104 can slide independently the time scale top in the screen lower part, respectively. The evaluation term of drawing will have been from 9:00 a.m. on May 30 till 11:00 p.m.

[0033] If an evaluation term is chosen, it will judge whether the following step 62 (drawing 3) has the performance criteria which is not "as a specification" within the selected evaluation term. As for this decision, it is desirable to be carried out by the performance analysis utility which resides in the computer station 12 permanently. Judgment of step 62 can be made by various technique. One technique is taking out the average of the sample of each performance criteria measured during which was specified] the evaluation. Next, it can judge whether as compared with each specification threshold, performance metrics are "as a specification" about those averages. A sample is combined among [other than the technique of taking the

average] two or more performance metrics, and complicated technique can be used rather than it filters them and it draws compound metrics.

[0034] The example of two metrics decisions is shown in following Table 2. Although the result of "SCSI queue" is determined from simple evaluation of the performance metrics "maximum SCSI queue length", the result of "bandwidth" is the example of the compound metrics which consist of the combination of the "host read block count" which is the sampled raw performance metrics, "the host write-in block count", and "activity time."

[0035]

[Table 2]

表 2				
名前	測定基準	仕様	仕様外	支援された決定
SCSI待ち行列	SCSI待ち行列最大値	0 - 1	≥ 3	4) ディスクの追加は助けになるか?
帯域幅	{(ホストが読み取ったブロック数+ホストが書き込んだブロック数)*512}÷活動時間	10 ⁶	10 ⁷	1) ボトルネックがあるか?

[0036] Other technique is checking the grade whose performance metrics' are out of specification ["out of specification"] "as a specification." This can be performed by setting up two or more thresholds or permission calculation values with which the different conformity grade is related about each performance metrics. The degree of nonconformity is judged based on the permission calculation value [which threshold or] the performance metrics sampled in the sampled performance metrics as compared with various thresholds fulfill. According to how much performance metrics are "out of specification", various conclusions about a data-storage system are drawn.

[0037] If performance metrics are based on the specification (in namely, the case of YES branching from the inquiry step 64 of drawing 3), the data-storage system will operate on optimum performance level. All the performance analysis utilities that reside in the computer station 12 permanently return the comment that it is favorable condition, and display it on the display monitor 18 (step 66). An user is notified of change being unnecessary by this feedback.

[0038] On the contrary, one or more performance metrics are not based on a specification, but, in a case (in namely, the case of "NO" branching from an inquiry of step 64) "out of specification", the proposal which is useful to a data-storage system raising performance to an user is shown (step 68). With which [of one or more performance metrics] is "out of specification", it is based how much it is "out of specification", and such a proposal is chosen from the list of proposals considered, when possible.

[0039] The one technique of leading to a suitable proposal is listing a predetermined proposal to a table and relating them with performance metrics. The example of the relation between the performance metrics decision "SCSI queue" of Table 2 and "bandwidth" (and other performance metrics), and a proposal is shown in Table 3.

[0040]

[Table 3]

表 3		
仕様外	仕様内	提案
SCSI待ち行列	最大ドライブ	ドライブにたいして作業負荷が重すぎます。条件を存続させるのなら、ドライブを追加してください。
SCSI待ち行列最大ドライブ		記憶システムに作業負荷が重すぎます。条件を存続させるのなら、記憶システムの追加か、より大きい記憶システムへのアップグレードを考慮してください。
帯域幅ホスト設定		パフォーマンスの最適化のためにホストデータの帯域幅の抑制が必要かもしれません。ホスト設定を調整してください。
帯域幅設定	ホスト設定	記憶システムはホストに依存するには帯域幅が不十分です。よりハイエンドの記憶システムにアップグレードしてください。

[0041] If the conditions of performance metrics are given and one or more proposals are chosen, the proposal will be displayed on an user by the display monitor 18 of a host computer (step 70). Example 106 of a performance planned proposal screen is shown in drawing 6 . This screen is related with the performance history illustrated to drawing 5 . The proposal "performance will improve if DRAM is increased" is displayed on this screen, and the indication which resulted in the proposal is listed further. Please observe that it is indicated that cache operating-condition metrics have become "out of specification", and the write-in cache has filled. Furthermore, note that DRAM with which it is equipped is not the maximum. Other indication shows the normal operation. If DRAM is added based on these results, the conclusion that performance improves is proposed. From this proposal, an user can perform decision based on an information and can decide now whether to add DRAM or to postpone an addition of DRAM.

[0042] When an user wants to evaluate the raw performance metrics which are the footing of the proposal of adding DRAM, the following step 72 (drawing 3) is displaying the performance metrics itself on the display monitor 18. The example 108 of a screen of performance metrics is shown in drawing 7 . "Relocation", a "work set", a "cache operating condition", "drive Q", and the performance metrics about "multi-read" are displayed. The same evaluation term from 9:00 a.m. on May 30 to 11:00 p.m. is divided by vertical lines 102 and 104.

[0043] This bar graph shows the processed data sample extracted during this period, in order to offer positive support of the

proposal which an user is shown. For example, please the cache operating-condition metrics in an evaluation term be comparatively high, therefore bring the result referred to as "out of specification", and note that a conclusion called an addition of DRAM is drawn. An user can evaluate various terms and can also investigate whether new mean metrics still "are not out of specification."

[0044] Please regard that the buttons 110, 112, and 114 for referring to quickly the performance history screen 100, the proposal screen 106, or the performance metrics screen 108, respectively are contained in each screen of the drawing 5 or the drawing 7. These three screens offer the abundant informations which are useful to pinpointing the field for the enhancement in a data-storage system of an user in performance. Moreover, these screens also serve as the help to which an user predicts the required storage capacity of the future and upgrades without retardation.

[0045] Since the means indicated on these specifications is what consists of the desirable mode which carries out this invention, this invention is not limited to the specific characteristic feature of having illustrated and explained. Therefore, a claim is asked for this invention in all modes or change within the limits of [suitable] a publication.

[0046] As mentioned above, although the example of this invention was explained in full detail, this inventions are hereafter enumerated for every embodiment.

(1) . The step which is the technique of pinpointing the field for the enhancement in performance in a data-storage system, and samples performance metrics working [data storage], The step which shows an user the performance history which shows how the data-storage system is operating during fixed, The step as which an user can choose the evaluation term of a performance history, The step which judges whether performance metrics have the level of the optimum performance of a data-storage system during the selected evaluation, How to have the step which provides an user with at least one proposal considered to be useful to raising the performance of a data-storage system when there are no performance metrics in the level of optimum performance.

(2) . Technique given in (1) to which a presentation step is characterized by an user having the step which enables it to choose the evaluation term of a performance history by enabling it to operate the term field where the selection step piled up the performance history on the performance history on the display screen including the step which carries out graphical display by the graphics.

(3) . Technique given in (1) to which a decision step is characterized by having the step which takes the average of the sample of the performance metrics measured during fixed, and asks for type average performance metrics.

(4) . Technique given in (1) characterized by having the step with which a decision step combines the sample of the performance metrics measured during fixed, and the step which calculates the type value which filters the combined sample and shows performance metrics.

(5) . The step which defines two or more performance metrics which are the technique of pinpointing the field for the enhancement in performance in a data-storage system, and can be measured working [a data-storage system], The step which a data-storage system presupposes is operated with optimum performance when the specification of performance metrics is offered and performance metrics follow the specification, The step which samples performance metrics working a data-storage system], The step which creates the performance history which shows how the data-storage system is operating during fixed based on the storage demand which the user issued, The step which shows an user a performance history, and the step as which an user enables it to choose the evaluation term of a performance history, The step which judges whether there are any performance metrics according to a specification during the selected evaluation, When one or more performance metrics do not follow a specification Technique characterized by having the step which offers the proposal considered to be useful to raising the performance of a data-storage system based on any of one or more performance metrics do not follow the specification.

(6) . Technique given in (5) to which it has the threshold of each performance metrics, and a decision step takes the average of each performance metrics in an evaluation term, and a specification is characterized by having a step [metrics / mean performance / each threshold].

(7) . Technique given in (5) to which a decision step is characterized by having the step which checks how much one or more performance metrics do not follow the specification.

(8) . Technique given in (5) characterized by for a specification having two or more thresholds of each performance criteria, and having the step to which the decision step of each performance metrics takes the average of the performance metrics in an evaluation term, a step [metrics / mean performance / two or more thresholds], and the step which computes the grade of a nonconformity based on any of two or more thresholds are over performance metrics.

(9) . Technique given in (5) to which it has the step to which a presentation step carries out graphical display of the performance history on the display screen, and a selection step piles up a term field by the graphic on a performance history, and is characterized by having the step which enables it to choose an evaluation term when an user operates a term field by the graphic.

(10) . Technique given in (5) to which it has the step which lists two or more predetermined proposals on the table which relates a proposal with performance metrics, and a proposal offer step is characterized by having the step which extracts one or more proposals from a table based on any of one or more performance metrics are not based on the specification.

[0047]

[Effect of the Invention] As mentioned above, the technique of this invention can pinpoint the potential field for the enhancement in performance, can tell an user about the performance below a canonical, and can generate the proposal which supports that an user corrects a problem and returns a system to optimum performance.

[Translation done.]

*** NOTICES ***

Japan Patent Office is not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] It is drawing showing the host-computer station connected to the data-storage system of this invention.

[Drawing 2] It is the block diagram of the data-storage system carried out as a hierarchy disk array.

[Drawing 3] It is the flow chart having shown how to pinpoint the field for the enhancement in a data-storage system in performance.

[Drawing 4] It is the ** type view showing the location which can measure the data flow and performance metrics in a data-storage system.

[Drawing 5] It is the display screen as which the performance history of a data-storage system was displayed.

[Drawing 6] It is the display screen as which the proposal about a performance plan was displayed.

[Drawing 7] It is the display screen as which the measured performance metrics were displayed.

[Description of Notations]

50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72: block

[Translation done.]

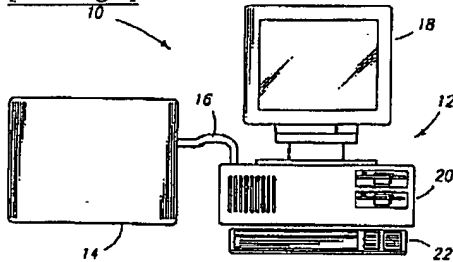
* NOTICES *

Japan Patent Office is not responsible for any damages caused by the use of this translation.

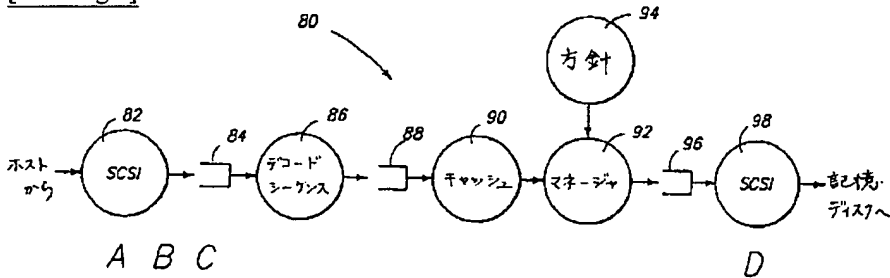
1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DRAWINGS

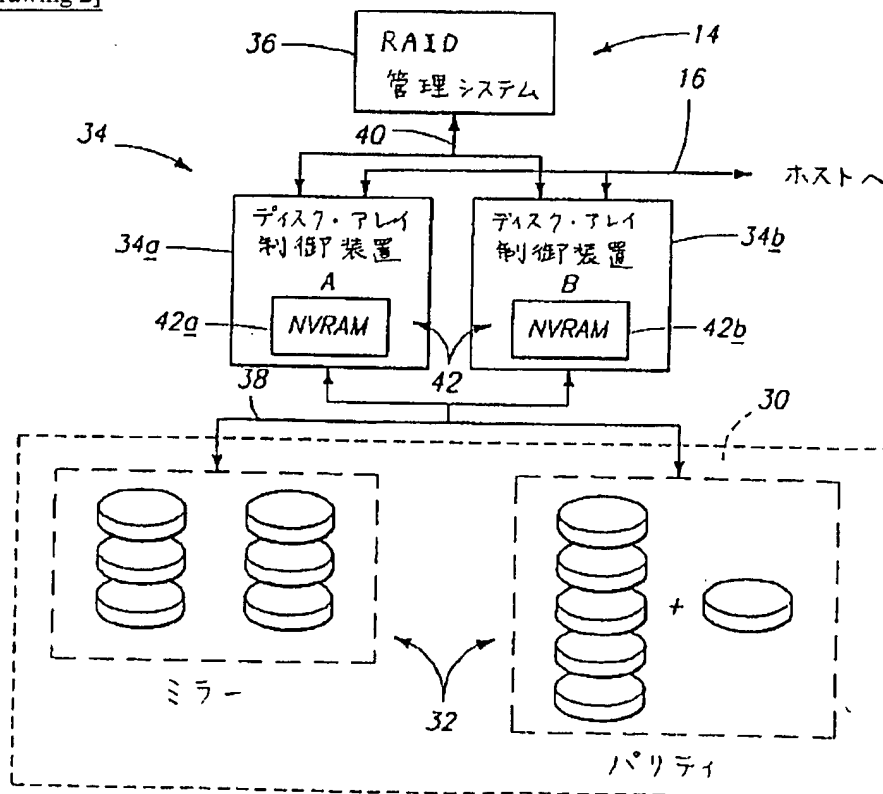
[Drawing 1]



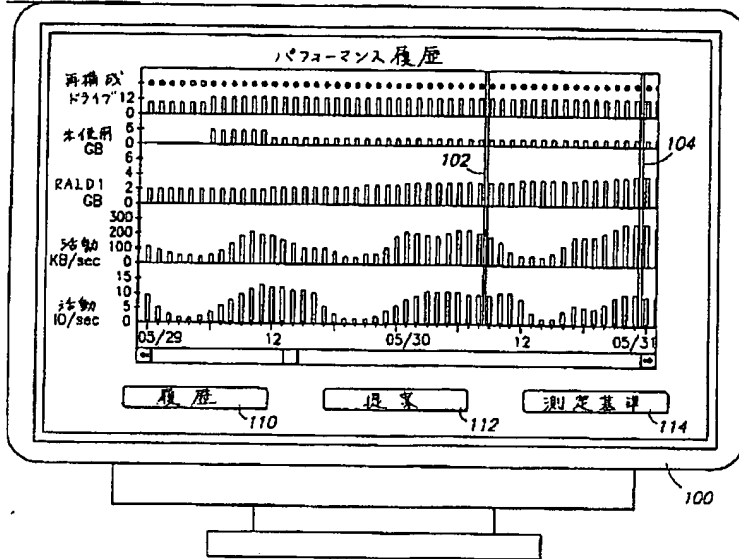
[Drawing 4]



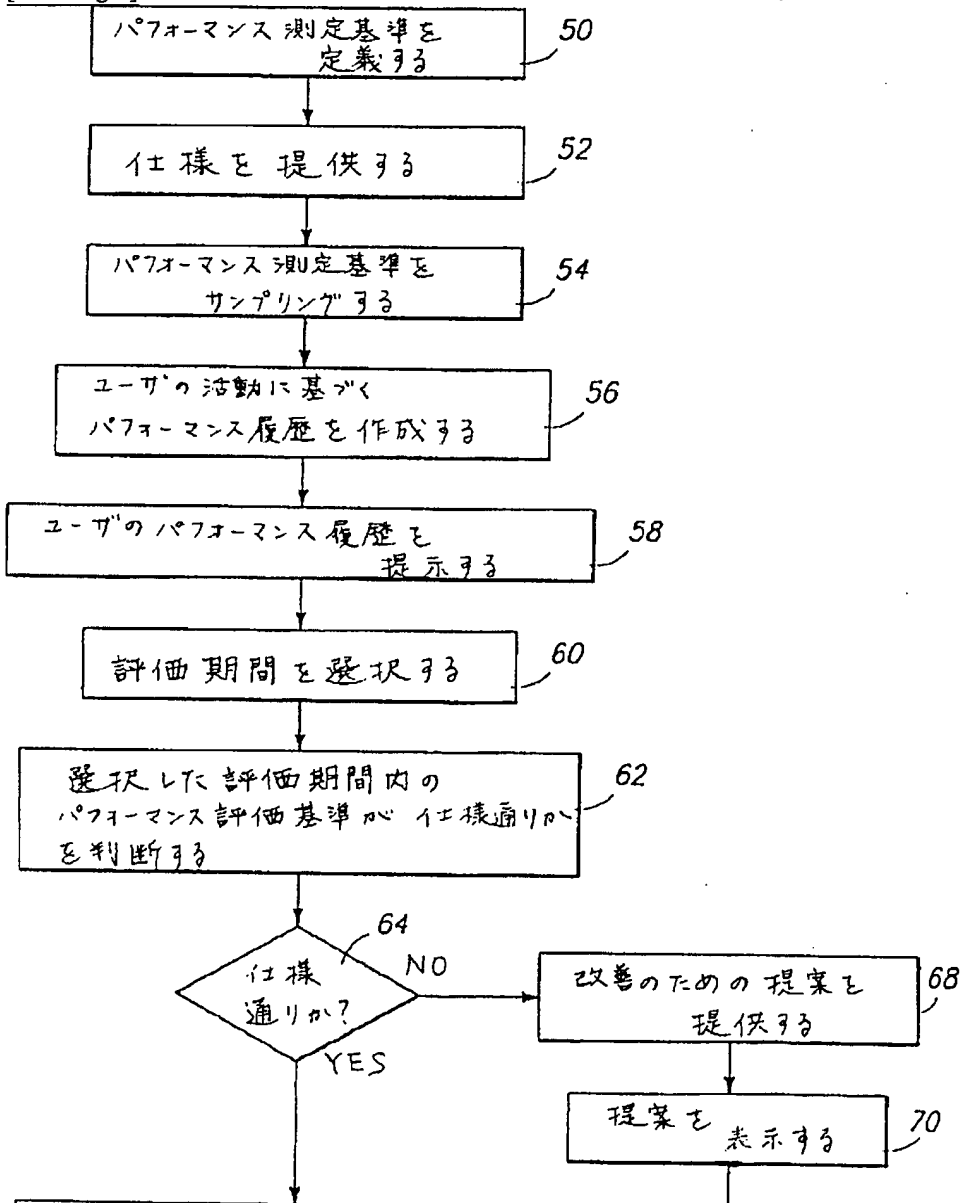
[Drawing 2]



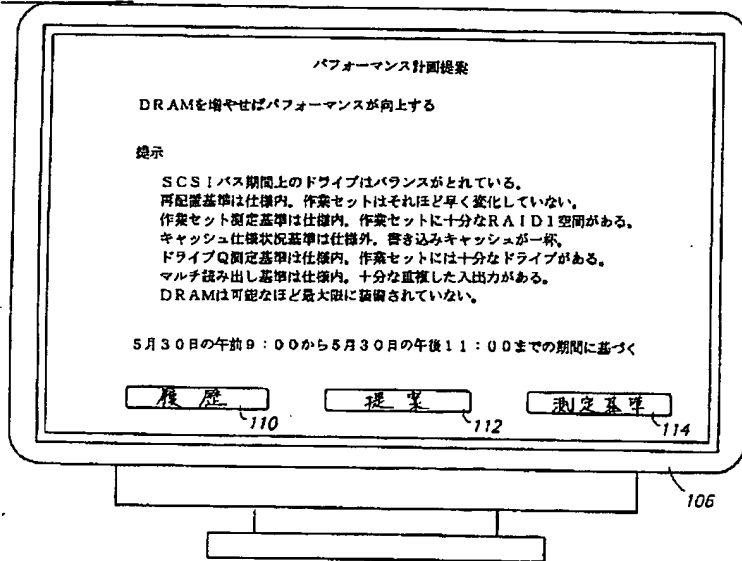
[Drawing 5]



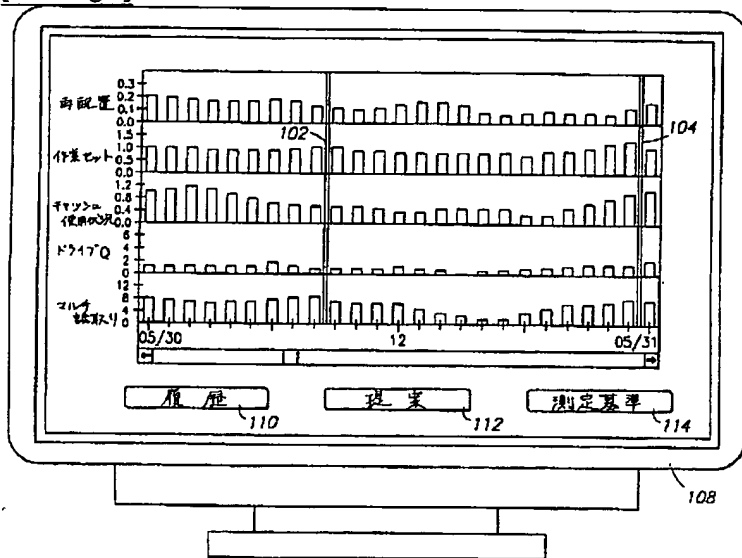
[Drawing 3]



[Drawing 6]



[Drawing 7]



[Translation done.]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-241227

(43) 公開日 平成8年(1996)9月17日

(51) Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 11/34		7313-5B	G 0 6 F 11/34	S
3/06	3 0 4		3/06	3 0 4 P

審査請求 未請求 請求項の数10 F D (全 12 頁)

(21) 出願番号 特願平7-328336

(22) 出願日 平成7年(1995)11月22日

(31) 優先権主張番号 343, 809

(32) 優先日 1994年11月22日

(33) 優先権主張国 米国 (US)

(71) 出願人 590000400

ヒューレット・パカード・カンパニー
アメリカ合衆国カリフォルニア州パロアル
ト ハノーバー・ストリート 3000

(72) 発明者 ダグラス・エル・ボイト

アメリカ合衆国アイダホ州ボイジー、エヌ
24ス 3030

(72) 発明者 マービン・ディー・ネルソン

アメリカ合衆国アイダホ州ボイジー、サン
フラワー・レイン 9055

(72) 発明者 ウェイド・エイ・ドルフィン

アメリカ合衆国アイダホ州ボイジー、ペト
リー 7179

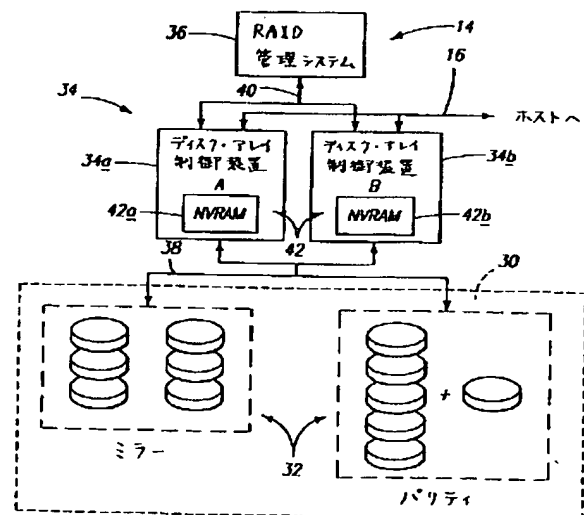
(74) 代理人 弁理士 上野 英夫

(54) 【発明の名称】 記憶システムのパフォーマンス向上特定方法

(57) 【要約】

【課題】 パフォーマンスに悪影響を及ぼす状況を監視または予測する目的で動作中のシステム・パフォーマンスを評価し、パフォーマンスが最適でないことが検出されたら、それをユーザに警告し、パフォーマンスを向上させるために講ずることができる処置について指示する。

【解決手段】 データ記憶システムの動作中に測定する複数のパフォーマンス測定基準を定義し(50)、その期待値の仕様も提供する(52)。システムの動作中にパフォーマンス測定基準をサンプリングし(56)、パフォーマンス履歴を作成し、システムが一定期間中にどのように動作していたかが示される(58)。評価期間を選択し、任意の測定基準が仕様に従っているかを判断する(62)。1つ以上のパフォーマンス測定基準が仕様に従っていない場合は、データ記憶システムのパフォーマンスを向上させるのに役立つ1つ以上の提案が与えられる(68)。



1

【特許請求の範囲】

【請求項1】データ記憶システムにおけるパフォーマンス向上のための領域を特定する方法であって、データ記憶装置の動作中にパフォーマンス測定基準をサンプリングするステップと、一定期間中にデータ記憶システムがどのように動作しているかを示すパフォーマンス履歴をユーザに提示するステップと、ユーザがパフォーマンス履歴の評価期間を選択することができるステップと、選択した評価期間中にパフォーマンス測定基準がデータ記憶システムの最適パフォーマンスのレベルになっているかどうかを判断するステップと、パフォーマンス測定基準が最適パフォーマンスのレベルにない場合、データ記憶システムのパフォーマンスを向上させるのに役立つと考えられる少なくとも1つの提案をユーザに提供するステップとを有する方法。

【請求項2】提示ステップが、表示画面上にパフォーマンス履歴をグラフィック表示するステップを含み、選択ステップが、パフォーマンス履歴上に重ね合わされた期間領域をグラフィックスで操作できるようにすることによってユーザがパフォーマンス履歴の評価期間を選択できるようにするステップを有することを特徴とする、請求項1に記載の方法。

【請求項3】判断ステップが、一定期間中に測定されたパフォーマンス測定基準のサンプルの平均値をとって典型平均パフォーマンス測定基準を求めるステップを有することを特徴とする、請求項1に記載の方法。

【請求項4】判断ステップが、一定期間中に測定されたパフォーマンス測定基準のサンプルを組み合わせるステップと、組み合わせたサンプルをフィルタリングしてパフォーマンス測定基準を示す典型値を求めるステップとを有することを特徴とする、請求項1に記載の方法。

【請求項5】データ記憶システムにおいてパフォーマンス向上のための領域を特定する方法であって、データ記憶システムの動作中に測定することができる複数のパフォーマンス測定基準を定義するステップと、パフォーマンス測定基準の仕様を提供し、パフォーマンス測定基準が仕様に従っている場合にはデータ記憶システムが最適パフォーマンスで動作しているとするステップと、データ記憶システムの動作中にパフォーマンス測定基準をサンプリングするステップと、ユーザが出した記憶要求に基づいて、一定期間中にデータ記憶システムがどのように動作しているかを示すパフォーマンス履歴を作成するステップと、ユーザにパフォーマンス履歴を提示するステップと、ユーザがパフォーマンス履歴の評価期間を選択できるよ

2

選択された評価期間中に仕様に従っていないパフォーマンス測定基準がないかどうかを判断するステップと、1つまたは複数のパフォーマンス測定基準が仕様に従っていない場合には、1つまたは複数のパフォーマンス測定基準のいずれが仕様に従っていないかに基づいて、データ記憶システムのパフォーマンスを向上させるのに役立つと考えられる提案を提供するステップとを有することを特徴とする方法。

【請求項6】仕様が個々のパフォーマンス測定基準のしきい値を有し、判断ステップが、評価期間中の個々のパフォーマンス測定基準の平均値をとって、平均パフォーマンス測定基準をそれぞれのしきい値と比較するステップを有することを特徴とする、請求項5に記載の方法。

【請求項7】判断ステップが、1つまたは複数のパフォーマンス測定基準がどの程度仕様に従っていないかを確認するステップを有することを特徴とする、請求項5に記載の方法。

【請求項8】仕様が個々のパフォーマンス基準の複数のしきい値を有し、個々のパフォーマンス測定基準の判断ステップが、評価期間中のパフォーマンス測定基準の平均値をとるステップと、

平均パフォーマンス測定基準を複数のしきい値と比較するステップと、複数のしきい値のいずれがパフォーマンス測定基準を超えているかに基づいて非準拠の程度を算出するステップとを有することを特徴とする、請求項5に記載の方法。

【請求項9】提示ステップが、表示画面上にパフォーマンス履歴をグラフィック表示するステップを有し、選択ステップが、パフォーマンス履歴の上に期間領域をグラフィックで重ね合わせて、ユーザが期間領域をグラフィックで操作することによって評価期間を選択できるようにするステップを有することを特徴とする、請求項5に記載の方法。

【請求項10】提案をパフォーマンス測定基準と関連づけるテーブルに、複数の所定の提案をリストするステップを有し、提案提供ステップが、1つまたは複数のパフォーマンス測定基準のいずれが仕様に準拠していないかに基づいて、テーブルから1つまたは複数の提案を抽出するステップを有することを特徴とする、請求項5に記載の方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、ディスク・アレイ・データ記憶システムなどのデータ記憶システムと、ユーザの要求に基づいてそのパフォーマンス向上技法を特定する方法とに関する。

【従来の技術】データ記憶システムのパフォーマンスは、その耐用期間中に必然的に変化する。パフォーマンスは、記憶システムで用いられている構成要素のタイプ、動作設定値、容量や入出力要求のようなユーザの記憶要件など、様々な要因の影響を受ける。構成要素が障害を起こしたり、使用率が過度または不十分になったりすると、パフォーマンスが悪影響を受けることがある。動作設定値がユーザの記憶構成にとって不適切な場合も、パフォーマンスが低下することがある。ユーザが、現行システムにとって過酷過ぎる記憶要求を出すと、記憶システムに過度の負荷がかかり、パフォーマンスが悪影響を受けることになる。

【0003】このようなパフォーマンスに影響を与える諸要因は、データ記憶システムの耐用期間中に変化する。さらに、このような要因は短期間で大幅に変動することがある。パフォーマンスが著しく変化したとき、ユーザはしばしば、パフォーマンスを向上させると予測または確信される、システムへの変更を加えようとする。しかし、そのような変更は、看取された問題に基づいて行われるが、その問題は実際にはより根本的な隠れた問題の徴候であることがある。このような情報に基づかない問題解決によって、ユーザは、満足なパフォーマンス向上が得られずに、不要なコストやシステム停止時間を生じさせることがある。たとえば、実際にはユーザの記憶要求の増加に対応する新たな方式でデータ記憶を再構成するだけでパフォーマンスが向上するにもかかわらず、ユーザはディスク・アレイにディスクをもう一つ追加すればパフォーマンスが向上すると考えることがある。

【0004】

【発明が解決しようとする課題】したがって、パフォーマンスに悪影響を及ぼす状況を監視または予測する目的で動作中のシステム・パフォーマンスを評価すれば有利なはずである。システムのパフォーマンスが最適でないことが検出されたら、それをユーザに警告し、パフォーマンスを向上させるために講ずることができる処置について指示することができよう。

【0005】

【課題を解決するための手段】本発明の1つの態様によれば、データ記憶システム内の改善すべき領域を特定する方法が提供される。データ記憶システムの動作中に測定することができる複数のパフォーマンス測定基準を定義する。パフォーマンス測定基準の例としては、ホストが読み取った合計ブロック数、ホストが書き込んだ合計ブロック数、ホストからコマンドが発行されていた時間の長さなどがある。パフォーマンス測定基準の期待値の仕様も提供する。パフォーマンス測定基準が仕様に従っている場合は、データ記憶システムは最適パフォーマンスで動作している。

動作中にパフォーマンス測定基準をサンプリングする。動作中に、パフォーマンス履歴が作成され、データ記憶システムが一定期間中にどのように動作していたかが示される。パフォーマンス履歴によって、1秒当たりの入出力数、メモリ・トランザクション数、使用済みおよび未使用のメモリ空間など、ユーザによって行われた活動と記憶要求が検査される。パフォーマンス履歴はユーザに提供され、一定期間中の活動と記憶要求をプロットしたグラフの形であることが好ましい。

10 【0007】パフォーマンス履歴を調べるとき、ユーザはより精密に精査するために評価期間を選択することができる。これは、パフォーマンス履歴上に期間領域をグラフィックを使用して重ね合わせ、その期間領域をユーザがグラフィックで操作して所望の時間枠に重ねることによって、評価期間を選択できるようにして実現することが好ましい。選択した評価期間について、任意のパフォーマンス測定基準が仕様に従っているかどうかを判断することができる。1つまたは複数のパフォーマンス測定基準が仕様に従っていない場合は、データ記憶システムのパフォーマンスを向上させるのに役立つ1つまたは複数の様々な提案が与えられる。提案は、どのパフォーマンス測定基準が仕様に従っていないかに基づいて選択される。

【0008】これらの提案の使用により、ユーザはデータ記憶システムのパフォーマンスを向上させる方法を、情報に基づいて決定することができる。また、本発明のパフォーマンス計画方法によって、ユーザは将来の必要記憶容量を予測し、それに対応するようにデータ記憶システムを適時的にアップグレードすることもできる。

30 【0009】

【実施例】本発明の方法について、複数のディスク・ドライブを使用するコンピュータ・ディスク・アレイ・データ記憶システムに関して説明する。しかし、本発明の原理は、オープンリール記憶システムや、テープ・バックアップ・システム、CD記憶システムなど他のタイプのデータ記憶システムにも適用可能である。

【0010】図1に、入出力インタフェース・バス16を介してデータ記憶装置14に接続されているホスト・コンピュータ端末またはステーション12を有するコンピュータ・システム10を示す。ホスト・コンピュータ・ステーション12は、表示モニタ18、中央演算処理装置(CPU)20、およびキーボード22を備えている。

40 【0011】図2に、階層ディスク・アレイ・データ記憶システムとして実施されたデータ記憶システム14の構成例を示す。記憶システムは、複数の記憶ディスク32を有するディスク・アレイ30、ディスク・アレイ30に接続され、記憶ディスク32との間のデータ伝送を調整するディスク・アレイ制御装置34、およびRAI

【0012】「RAID」(独立ディスクの冗長アレイ)という用語は、物理記憶容量の一部を使用して、記憶容量の残りの部分に記憶されたユーザ・データに関する冗長情報を記憶する、ディスク・アレイを意味する。アレイの構成ディスクの1つまたはアクセス・バスに障害が発生した場合、この冗長情報によってユーザ・データを再生成することができる。RAIDシステムの詳細については、アメリカ合衆国ミネソタ州リノレイクのRAID Advisory Boardから1993年6月9日に刊行された「The RAIDBook: A Source Book for RAID Technology」という書物に記載されている。

【0013】ディスク・アレイ制御装置34は、小型コンピュータ・システム・インタフェース(SCSI)など1つまたは複数のインタフェース・バス38を介してディスク・アレイ30に接続されている。RAID管理システム36は、インタフェース・プロトコル40を介してディスク・アレイ制御装置34に作用を及ぼすようにして接続されている。制御装置34は、入出力インタフェース・バス16を介してホスト・コンピュータに接続されている。RAID管理システム36は、別個の構成要素として実施するか、ディスク・アレイ制御装置34内またはホスト・コンピュータ内に構成することができ、ディスク記憶および信頼性レベルを制御し、様々な信頼性記憶レベル間でデータを伝送するための、データ管理手段を提供する。これらの信頼性記憶レベルは、ミラーまたはパリティ冗長性レベルであることが好ましいが、冗長性がまったくない信頼性記憶レベルを含めることもできる。

【0014】ディスク・アレイ制御装置34は、ディスク・アレイ制御装置A 34aとディスク・アレイ制御装置B 34bとを備えた二重制御装置として実施される。二重制御装置34aおよび34bは、一方の制御装置が動作不能になった場合に途切れのないバックアップと冗長性を提供することによって、信頼性を強化する。

【0015】記憶ディスク32はたとえば1ないし3ギガバイトの容量を有する。記憶ディスクは、SCSIバス38とインタフェースさせる機械ベイにそれぞれ独立して接続したり切り離したりすることができる。1つの実施態様では、12個の活動機械ベイを備えたデータ記憶システムが設計される。4本のSCSIバスを使用して、これらのベイとディスク・アレイ制御装置34とをインタフェースさせることができる(すなわち機械ベイ3個について1本のバスを使用する)。活動ベイがすべて装荷されると、このデータ記憶システムの例示の複合容量は12ないし36ギガバイトになる。

【0016】階層ディスク・アレイ30は、その物理記憶空間と1つまたは複数の仮想記憶空間とを含むそれぞれ別々の記憶空間として特徴づけることができる。記憶装置のこれらのさまざまなビューは、マッピング技法に

物理記憶空間を、様々なデータ信頼性レベルに従って記憶域を表すRAIDレベルの仮想記憶空間にマップすることができる。たとえば、RAIDレベルの仮想記憶空間内の一部の領域を、ミラーまたはRAIDレベル1などの第1の信頼性記憶レベルのために割り振り、他の領域をパリティまたはRAIDレベル5などの第2の信頼性記憶レベルのために割り振ることができる。RAIDレベルの仮想ビューは、連続アドレス指定可能記憶空間を示す第2のアプリケーション・レベルの仮想記憶空間にマップすることができる。記憶空間の物理構成とRAIDビューは、ユーザに対して示されるアプリケーション・ビューからは隠されている。

【0017】様々な記憶空間を相互にマップするために用いられる仮想マッピング情報の持続記憶のために、メモリ・マップ記憶42が用意されている。メモリ・マッピング情報は、様々なビュー間の様々なマッピング構成が変化するに従って、制御装置またはRAID管理システムによって絶えずまたは定期的に更新することができる。メモリ・マップ記憶42は、バッテリー・バックアップされたRAM(ランダム・アクセス・メモリ)などの2つの不揮発性RAM(NVRAM)42aおよび42bとして実施され、それぞれの制御装置34aおよび34b内に配置される。二重NVRAM42aおよび42bは、メモリ・マッピング情報の冗長記憶のために設けられている。

【0018】記憶ディスクは、ホスト・コンピュータがデータ記憶に使用することができる物理記憶容量を提供する。この容量を有効に管理して、動作効率とユーザの記憶要求の満足という両方の観点から、パフォーマンスを最適化することが望ましい。本発明のパフォーマンス管理技法の基礎は、データ記憶システムがどれだけ効果的に動作しているかを絶えず評価するという困難で正確な問いに対する回答を見い出すことにある。以下の7項の質問で、1組のパフォーマンス向上意思決定目標の例を示す。

【0019】1. データ記憶システムはパフォーマンスの妨害になっているか。

2. 現行のシステム構成設定値によって最適パフォーマンスが可能か。設定値の例としては、SCSI転送速度や論理装置のサイズなどがある。

3. 4本のSCSIチャネル38に接続されている記憶ディスクのバランスを取り直すことによってパフォーマンスが向上するか。

4. 記憶ディスクを大容量化、高速化、または追加すればパフォーマンスが大幅に向上するか。

5. ディスク・アレイ制御装置34内のRAMを追加すればパフォーマンスが大幅に向上するか。

6. データ記憶システム14を追加すればパフォーマンスが大幅に向上するか。

記憶ディスク32にアップグレードすれば、パフォーマンスが向上するか。

【0020】第1項に挙げた意思決定の質問は、原因は突きとめられなくても何かがシステムに悪影響を及ぼしていることを示しており、質問2ないし5項は、既存のデータ記憶システムに変更を加える必要があることを示している。質問6および7項は、1つのデータ記憶システムの容量を全部使い果たした後の変更など、追加の構成要素を必要とする変更を要する。

【0021】ある程度予測を必要とするため、パフォーマンス向上に関する質問に絶対的な正確さをもって回答するのは難しい。しかし、パフォーマンス・パラメータの完全な数量的見積もりには、高い記憶コストと分析処理コストが必要である。したがって、本発明の方法は、データ記憶システムの既存の構成要素を利用することに焦点を合わせる。使用状況データは、パフォーマンスの問題を特定するのに役立ち、現在の構成で有効に動作している構成要素と劣っている構成要素を特定するのに役立つ。特定後の基本パフォーマンス向上戦略は、すでに仕様範囲内で動作している構成要素の変更ではなく、現在の問題を解決する変更へ投資することである。

【0022】本発明の方法は、パフォーマンス向上のための潜在的領域を特定する。これらの方法は、標準以下のパフォーマンスをユーザに知らせ、ユーザが問題を修正してシステムを最適パフォーマンスに戻すのを支援する提案を生成する。

【0023】図3に、パフォーマンス向上のための領域を特定する方法の好ましいステップを示す。最初のステップ50は、データ記憶システムの動作中に測定することができる1つまたは複数のパフォーマンス測定基準を定義することである。パフォーマンス基準の例としては、ホストが読み取る合計ブロック数、ホストが書き込む合計ブロック数、ホストからコマンドが送られていた時間の長さ、装置当たりのSCSI待ち行列長などがある。パフォーマンス測定基準の標準を設定する仕様も提供される(ステップ52)。データ記憶システムは、パフォーマンス測定基準がこの仕様に従っているとき、言い換えると、パフォーマンス測定基準が「仕様通り」で

あるときに、最適パフォーマンスで動作している。仕様によって設定する標準は、様々なパフォーマンス測定基準のしきい値による受容算定数値または許容算定数値の形態をとることができる。

【0024】パフォーマンス測定基準およびそれに対応する仕様を定義した後、データ記憶システムにパフォーマンス測定基準を監視する手段を設ける。このような手段は、動作中の構成要素使用状況データを抽出するディスク・アレイ制御装置34内のファームウェアとして実施することが好ましい。

【0025】パフォーマンス測定基準を導入した後、次のステップ54は、データ記憶装置の動作中のパフォーマンス測定基準をサンプリングすることである。このサンプリングは、データ記憶システム14またはホスト・コンピュータ・ステーション12で行う。図4に、パフォーマンス測定基準をサンプリングすることができる点を示した制御装置ファームウェアの模式例80を示す。円は、入出力処理の記憶を表し、3辺を持つボックスは待ち行列を表す。この模式図は、ホスト・コンピュータから記憶ディスクまでの情報の流れを示している。ホストからの要求はまずフロント・エンドのSCSI 82に流れ、次にホストの待ち行列84に一時的に入れられる。それ以降の処理のために、デコード86が要求をより小さいコマンドに細分化する。コマンドは、キャッシュ90の受入れ準備が整うまでキャッシュ待ち行列88に入れられる。階層および冗長性管理に関するRAID方針94に応じて、ワークロードが階層マネージャ92で変換される。流れは、バック・エンドSCSI 98に送られる項目が一時的に入れられる、SCSI待ち行列96に続く。

【0026】パフォーマンス測定基準は、この模式上の様々な点で抽出することができる。たとえば、ホストが読み取った合計ブロック数の測定基準は、フロント・エンドSCSI 82のA点で抽出される。以下の表1に、測定基準例、パフォーマンス測定基準の抽出点、およびパフォーマンス測定基準に影響される決定を示す。

【0027】

【表1】

表1		
測定基準	サンプリング点	支援された決定
活動時間—ホストから>0命令が顕著な間の時間の合計	A	1) ボトルネックがあるか？ 7) アップグレードでパフォーマンス改善できるか？
SCSI待ち行列最大値—装置当りの最大のSCSI待ち行列サイズ	D	3) ディスクのバランスを取り直すか？ 4) ディスクの追加は助けになるか？
ホスト読取りブロッカー—ホストにより読み取られたブロックの数の合計	B	1) ボトルネックがあるか？ 5) 制御RAMがもっと必要か？ 7) アップグレードでパフォーマンスを改善できるか？
ホスト書き込みブロック数—ホストにより書き込まれたブロックの数の合計	B	1) ボトルネックがあるか？ 5) 制御RAMがもっと必要か？ 7) アップグレードでパフォーマンスを改善できるか？

【0028】パフォーマンス測定基準は、データ記憶システム内の任意の場所で測定することができる。たとえば、C点ではホスト待ち行列の最大長に関するパフォーマンス測定基準のサンプルを抽出することができる。

【0029】図3に戻り、ステップ56で、ユーザの記憶活動に基づくパフォーマンス履歴が作成される。パフォーマンス履歴には、使用可能空間の量、消費された空間の量、入出力要求、ミラーおよびパリティ・タイプ記憶の量などのパラメータなど、データ記憶システムの使用を反映するパラメータが含まれる。パフォーマンス履歴によって、データ記憶システムが一定期間中にどのように動作していたかが明らかになる。ステップ56でのパフォーマンス履歴の作成は、データ記憶システム14またはホスト・コンピュータ・ステーション12で行うことができる。

【0030】次に、ステップ58でパフォーマンス履歴がホスト・コンピュータ・ステーション12でユーザに提示される。パフォーマンス履歴は図1の表示画面18にグラフで表示されることが好ましい。

【0031】図5に、パフォーマンス履歴の表示画面例100を示す。具体的には、3日間にわたるデータ記憶システムの未使用メモリ容量（ギガバイト単位で測定）、RAIDレベル1タイプの記憶量（ギガバイト単位で測定）、入出力アクセス活動（1秒当たりのキロバイト数と1秒当たりの入出力数の両方で測定）に関するパフォーマンスが示される。パフォーマンス履歴では、午前6時から午後12時までの日中かなりの活動が示されていることに注目されたい。表示期間はユーザが調整することができる。

【0032】ここで、ユーザが1つの特定の期間を評価したいとする。たとえば、ユーザはパフォーマンス履歴内の異常な瞬間や傾向を観察して、その期間をさらに綿密に調べたい場合がある。また別の例では、ユーザが比較的正常な期間を分析しているときに、非典型的な使用*

* 状況の期間を表示しないようにしたい場合がある。どちらの場合も、ユーザは、パフォーマンス履歴グラフ上に重ね合わされた垂直線102および104で区切られた期間領域をグラフィックで操作して評価期間を選択するだけで済む（図3のステップ60）。各境界線102、104は、画面下部にあるタイム・スケール上をそれぞれ独立してスライドすることができる。図の評価期間は5月30日の午前9時から午後11時までとなっている。

【0033】評価期間を選択したら、次のステップ62（図3）は、選択した評価期間内に「仕様通り」でないパフォーマンス基準があるかどうかを判断する。この判断は、コンピュータ・ステーション12に常駐するパフォーマンス分析ユーティリティによって行われることが好ましい。ステップ62の判断は様々な方法で行うことができる。1つの手法は、指定した評価期間中に測定された各パフォーマンス基準のサンプルの平均値を出すことである。次にそれらの平均値をそれぞれの仕様しきい値と比較して、パフォーマンス測定基準が「仕様通り」であるかどうかを判断することができる。平均値をとる方法の他に、複数のパフォーマンス測定基準間でサンプルを結合し、それらをフィルタリングして複合測定基準を導き出すというより複雑な技法を用いることができる。

【0034】以下の表2に、2つの測定基準決定の例を示す。「SCSI待ち行列」の結果は「最大SCSI待ち行列長」というパフォーマンス測定基準の単純な評価から決定されるが、「帯域幅」の結果は、サンプリングされた生パフォーマンス測定基準である「ホスト読取りブロック数」、「ホスト書き込みブロック数」および「活動時間」の組み合わせから成る複合測定基準の例である。

【0035】

【表2】

1 1

1 2

表2				
名前	測定基準	仕様	仕様外	支援された決定
SCSI待ち行列	SCSI待ち行列最大値	0-1	≥3	4) ディスクの追加は助けになるか?
帯域幅	{(ホストが読み取ったブロック数+ホストが書き込んだブロック数)*512}÷活動時間	10 ⁶	10 ⁷	1) ボトルネックがあるか?

【0036】他の手法は、パフォーマンス測定基準が「仕様通り」または「仕様外」である程度を確認することである。これは、各パフォーマンス測定基準について、異なる準拠程度が関連づけられている複数のしきい値または許容算定値を設定することによって行うことができる。サンプリングされたパフォーマンス測定基準を様々なしきい値と比較し、サンプリングされたパフォーマンス測定基準がどのしきい値または許容算定値を満たしているかに基づいて、非準拠度を判断する。パフォーマンス測定基準がどの程度「仕様外」であるかに応じて、データ記憶システムに関する様々な結論が導き出される。

【0037】パフォーマンス測定基準が仕様に準拠していれば（すなわち、図3の問い合わせステップ64からのYES分岐の場合）、データ記憶システムは最適パフォーマンス・レベルで動作している。コンピュータ・ステーション12に常駐するパフォーマンス分析ユーティリティは、すべて順調であるというコメントを返して表示モニタ18に表示する（ステップ66）。このフィー*

* ドバックにより、変更が不要であることがユーザに通知される。

10 【0038】反対に、1つまたは複数のパフォーマンス測定基準が仕様に準拠しておらず、「仕様外」の場合（すなわちステップ64の問い合わせからの「NO」分岐の場合）、データ記憶システムはユーザに、パフォーマンスを向上させるのに役立つ提案を提示する（ステップ68）。このような提案は、1つまたは複数のパフォーマンス測定基準のどれが「仕様外」であるかということと、可能な場合にはどの程度「仕様外」であるかに基づいて、考えられる提案のリストから選択される。

20 【0039】適切な提案に導く1つの方法は、所定の提案を表にリストし、それらをパフォーマンス測定基準と関連づけることである。表3に、表2のパフォーマンス測定基準決定「SCSI待ち行列」および「帯域幅」（およびその他のパフォーマンス測定基準）と提案との関係の例を示す。

【0040】

【表3】

表3		
仕様外	仕様内	提案
SCSI待ち行列	最大ドライブ	ドライブにたいして作業負荷が重すぎます。条件を存続させるのなら、ドライブを追加してください。
SCSI待ち行列最大ドライブ		記憶システムに作業負荷が重すぎます。条件を存続させるのなら、記憶システムの追加か、より大きい記憶システムへのアップグレードを考慮してください。
帯域幅ホスト設定		パフォーマンスの最適化のためにホストデータの帯域幅の抑制が必要かもしれません。ホスト設定を調整してください。
帯域幅	ホスト設定	記憶システムはホストに依存するには帯域幅が不十分です。よりハイエンドの記憶システムにアップグレードしてください。

【0041】パフォーマンス測定基準の条件が与えられ、1つまたは複数の提案が選択されると、その提案がホスト・コンピュータの表示モニタ18によってユーザに表示される（ステップ70）。パフォーマンス計画提案画面の例106を図6に示す。この画面は、図5に図示したパフォーマンス履歴に関するものである。この画面には、「DRAMを増やせばパフォーマンスが向上する」という提案が表示され、さらに、その提案に至った徴候がリストされる。キャッシュ使用状況測定基準が「仕様外」になっており、書き込みキャッシュがいっぱいになっていると示されていることに注目されたい。さらに、装着されているDRAMが最大限でないことに注意

※これらの結果に基づいて、DRAMを追加すればパフォーマンスが向上するという結論が提案されている。この提案から、ユーザは情報に基づく決定を行い、DRAMを追加するかDRAMの追加を見合わせるかを決めることができるようになる。

40 【0042】ユーザが、DRAMを追加するという提案の基礎になっている生パフォーマンス測定基準を評価したい場合、次のステップ72（図3）は表示モニタ18にパフォーマンス測定基準そのものを表示することである。パフォーマンス測定基準の画面例108を図7に示す。「再配置」、「作業セット」「キャッシュ使用状況」、「ドライブQ」、および「マルチ読取り」に関する

02および104によって、5月30日の午前9時から午後11時までの同じ評価期間が区切られている。

【0043】この棒グラフは、ユーザに提示される提案の積極的支援を提供するためにこの期間中に抽出された処理済みデータ・サンプルを示している。たとえば、評価期間中のキャッシュ使用状況測定基準は比較的高く、したがって「仕様外」という結果になり、DRAMの追加という結論が導き出されることに注意されたい。ユーザは様々な期間を評価して、新しい平均測定基準が依然として「仕様外」になっていないかどうかを調べることができる。

【0044】図5ないし図7の各画面には、それぞれパフォーマンス履歴画面100、提案画面106、またはパフォーマンス測定基準画面108を迅速に参照するためのボタン110、112、および114が含まれていることに留意されたい。この3つの画面は、ユーザがデータ記憶システム内のパフォーマンス向上のための領域を特定するのに役立つ豊富な情報を提供する。また、これらの画面は、ユーザが将来の必要記憶容量を予測して遅滞なくアップグレードを行う手助けともなる。

【0045】本明細書で開示した手段は本発明を実施する好ましい態様から成るものであるため、本発明は図示および説明した特定の特徴に限定されるものではない。したがって本発明は、特許請求の範囲に記載の適切な範囲内のあらゆる態様または変更において請求されるものである。

【0046】以上、本発明の実施例について詳述したが、以下、本発明を各実施態様毎に列挙する。

(1)．データ記憶システムにおけるパフォーマンス向上のための領域を特定する方法であって、データ記憶装置の動作中にパフォーマンス測定基準をサンプリングするステップと、一定期間中にデータ記憶システムがどのように動作しているかを示すパフォーマンス履歴をユーザに提示するステップと、ユーザがパフォーマンス履歴の評価期間を選択することができるステップと、選択した評価期間中にパフォーマンス測定基準がデータ記憶システムの最適パフォーマンスのレベルになっているかどうかを判断するステップと、パフォーマンス測定基準が最適パフォーマンスのレベルにない場合、データ記憶システムのパフォーマンスを向上させるのに役立つと考えられる少なくとも1つの提案をユーザに提供するステップとを有する方法。

(2)．提示ステップが、表示画面上にパフォーマンス履歴をグラフィック表示するステップを含み、選択ステップが、パフォーマンス履歴上に重ね合わされた期間領域をグラフィックスで操作できるようにすることによってユーザがパフォーマンス履歴の評価期間を選択できるようにするステップを有することを特徴とする、

(1)に記載の方法。

パフォーマンス測定基準のサンプルの平均値をとって典型平均パフォーマンス測定基準を求めるステップを有することを特徴とする、(1)に記載の方法。

(4)．判断ステップが、一定期間中に測定されたパフォーマンス測定基準のサンプルを組み合わせるステップと、組み合わせたサンプルをフィルタリングしてパフォーマンス測定基準を示す典型値を求めるステップとを有することを特徴とする、(1)に記載の方法。

(5)．データ記憶システムにおいてパフォーマンス向上のための領域を特定する方法であって、データ記憶システムの動作中に測定することができる複数のパフォーマンス測定基準を定義するステップと、パフォーマンス測定基準の仕様を提供し、パフォーマンス測定基準が仕様に従っている場合にはデータ記憶システムが最適パフォーマンスで動作しているとするステップと、データ記憶システムの動作中にパフォーマンス測定基準をサンプリングするステップと、ユーザが出した記憶要求に基づいて、一定期間中にデータ記憶システムがどのように動作しているかを示すパフォーマンス履歴を作成するステップと、ユーザにパフォーマンス履歴を提示するステップと、ユーザがパフォーマンス履歴の評価期間を選択できるようにするステップと、選択された評価期間中に仕様に従っていないパフォーマンス測定基準がないかどうかを判断するステップと、1つまたは複数のパフォーマンス測定基準が仕様に従っていない場合には、1つまたは複数のパフォーマンス測定基準のいずれが仕様に従っていないかに基づいて、データ記憶システムのパフォーマンスを向上させるのに役立つと考えられる提案を提供するステップとを有することを特徴とする方法。

(6)．仕様が個々のパフォーマンス測定基準のしきい値を有し、判断ステップが、評価期間中の個々のパフォーマンス測定基準の平均値をとって、平均パフォーマンス測定基準をそれぞれのしきい値と比較するステップを有することを特徴とする、(5)に記載の方法。

(7)．判断ステップが、1つまたは複数のパフォーマンス測定基準がどの程度仕様に従っていないかを確認するステップを有することを特徴とする、(5)に記載の方法。

(8)．仕様が個々のパフォーマンス基準の複数のしきい値を有し、個々のパフォーマンス測定基準の判断ステップが、評価期間中のパフォーマンス測定基準の平均値をとるステップと、平均パフォーマンス測定基準を複数のしきい値と比較するステップと、複数のしきい値のいずれがパフォーマンス測定基準を超えているかに基づいて非準拠の程度を算出するステップとを有することを特徴とする、(5)に記載の方法。

(9)．提示ステップが、表示画面上にパフォーマンス履歴をグラフィック表示するステップを有し、選択ステップが、パフォーマンス履歴の上に期間領域をグラフ

15

クで操作することによって評価期間を選択できるようにするステップを有することを特徴とする、(5)に記載の方法。

(10)．提案をパフォーマンス測定基準と関連づけるテーブルに、複数の所定の提案をリストするステップを有し、提案提供ステップが、1つまたは複数のパフォーマンス測定基準のいずれが仕様に準拠していないかに基づいて、テーブルから1つまたは複数の提案を抽出するステップを有することを特徴とする、(5)に記載の方法。

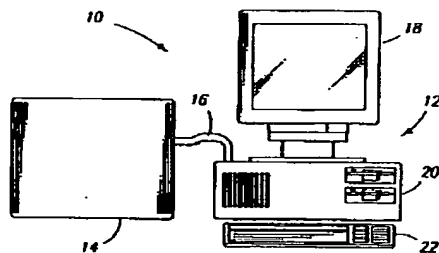
【0047】

【発明の効果】以上のように、本発明の方法は、パフォーマンス向上のための潜在的領域を特定し、標準以下のパフォーマンスをユーザに知らせ、ユーザが問題を修正してシステムを最適パフォーマンスに戻すのを支援する提案を生成することができる。

【図面の簡単な説明】

【図1】本発明のデータ記憶システムに接続されたホス

【図1】



16

ト・コンピュータ・ステーションを示す図である。

【図2】階層ディスク・アレイとして実施されたデータ記憶システムのブロック図である。

【図3】データ記憶システム内のパフォーマンス向上のための領域を特定する方法を示した流れ図である。

【図4】データ記憶システム内のデータ・フロー、およびパフォーマンス測定基準が測定できる場所を示す模式図である。

【図5】データ記憶システムのパフォーマンス履歴が表示された表示画面である。

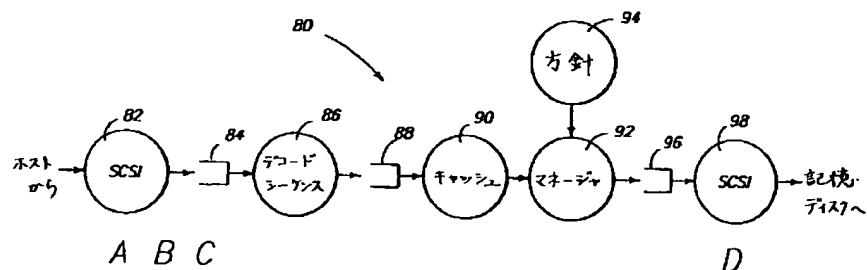
【図6】パフォーマンス計画に関する提案が表示された表示画面である。

【図7】測定されたパフォーマンス測定基準が表示された表示画面である。

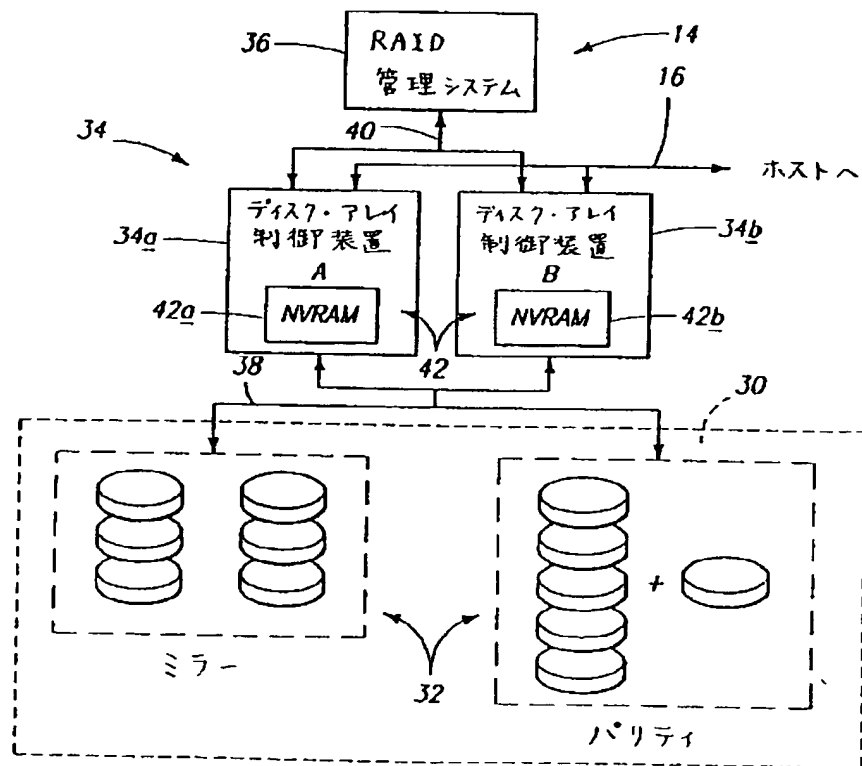
【符号の説明】

50、52、54、56、58、60、62、64、66、68、70、72：ブロック

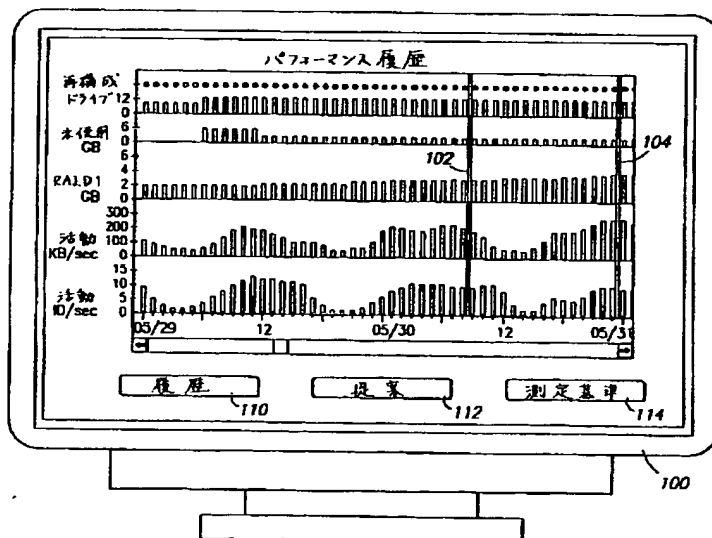
【図4】



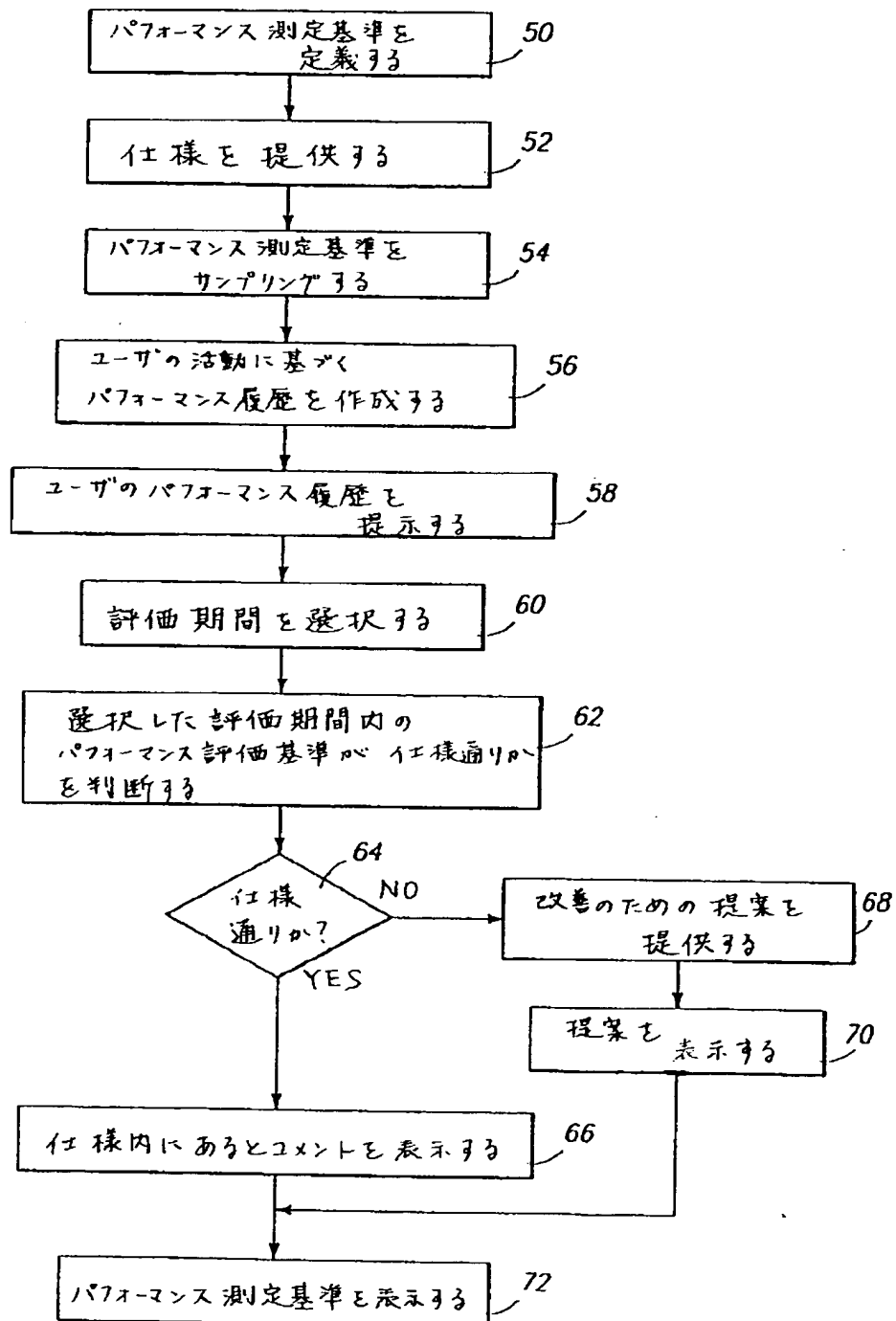
【図2】



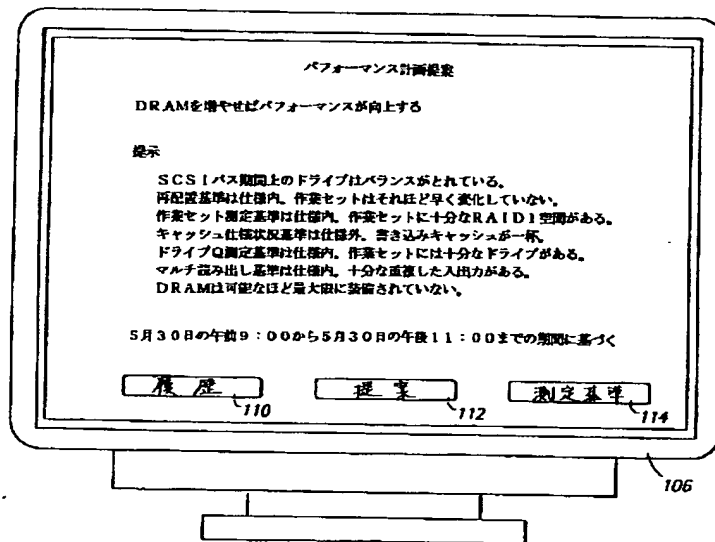
【図5】



【図3】



【図6】



【図7】

